

# Variable-Length Channel Quantizers for Maximum Diversity and Array Gains

Erdem Koyuncu and Hamid Jafarkhani

Center for Pervasive Communications and Computing, University of California, Irvine.

**Abstract**—We consider a  $t \times 1$  multiple-antenna fading channel with quantized channel state information at the transmitter (CSIT). Our goal is to maximize the diversity and array gains that are associated with the symbol error rate (SER) performance of the system. It is well-known that for both beamforming and precoding strategies, finite-rate fixed-length quantizers (FLQs) cannot achieve the full-CSIT diversity and array gains. In this work, for any function  $f(P) \in \omega(1)$ , we construct variable-length quantizers (VLQs) that can achieve these full-CSIT gains with rates  $1 + (f(P) \log P)/P$  and  $1 + f(P)/P^t$  for the beamforming and precoding strategies, respectively, where  $P$  is the power constraint of the transmitter. We also show that these rates are the best possible up to  $o(1)$  multipliers in their  $P$ -dependent terms. In particular, although the full-CSIT SER is not achievable at any (even infinite) feedback rate, the full-CSIT diversity and array gains can be achieved with a feedback rate of 1 bit per channel state asymptotically.

## I. INTRODUCTION

The performance of a multiple antenna communication system can be greatly improved by making the channel state information (CSI) available at the transmitter and/or the receiver. Typically, the receiver can acquire the CSI by training. Obtaining CSI at the transmitter (CSIT) is however more difficult and generally requires receiver's feedback. In this context, a channel quantizer specifies (i) for each channel state, the sequence of feedback bits to be fed back by the receiver; and (ii) for each such sequence, the transmission codeword (e.g. a beamforming vector) to be employed by the transmitter. The goal is then to design an optimal quantizer with respect to a specific performance measure (such as the symbol error rate (SER)) subject to the rate constraint of the feedback link.

An overview on channel quantizers for multiple antenna systems can be found in [1]. In particular, limited feedback beamforming has been extensively studied via Grassmannian line packings [2]–[4], quantizer design algorithms [5]–[8], high resolution methods [9], random vector quantizers [10], [11], and several other techniques [12]–[14]. Beamforming is a special case of the more general, albeit more complex transmission strategy called precoding, whose performance with quantized feedback has been studied in [15]–[20].

Most of the previous work on CSI feedback employ fixed-length quantizers (FLQs), in which the number of feedback bits per channel state is a fixed integer. In general, different binary codewords of different lengths can be fed back for different channel states, resulting in a variable-length quantizer (VLQ). An FLQ is clearly a special case of a VLQ, and we thus expect VLQs to provide a better performance than FLQs.

In this work, we focus on the VLQ design problem for a  $t \times 1$  multiple-input single-output (MISO) system. Our goal is

to maximize the diversity and array gains corresponding to the symbol error rate (SER) performance of the system (The reason why we will not focus on the SER itself will be explained later on.). We assume a quasi-static block fading channel model in which the channel realizations vary independently from one fading block to another while within each block they remain constant. We also assume that the receiver has full CSI, while the transmitter has only partial CSI provided by the receiver via error-free and delay-free feedback channels. The partial CSI is in the form of quantized instantaneous CSI provided by a VLQ. We design VLQs for beamforming and precoding strategies. For the latter strategy, we focus on linear precoding of complex orthogonal space-time block codes [21].

As we have shown previously, for a wide variety of communication problems, VLQs outperform FLQs by a significant margin [22]–[24]. In [22], we have considered the VLQ design problem for the outage probability performance measure. Compared to [22], although the general idea behind designing good VLQs will remain the same, the specific methods of [22] are not directly applicable due to fundamentally different distortion functions. Also as a result of this difference, we obtain completely different results and reach different conclusions.

In [23], [24], we have designed SER-optimizing distributed VLQs for beamforming in networks with multiple receivers. In this paper, the simpler non-distributed (point-to-point) nature of quantization allows us to prove much stronger achievability results. Here, we also prove converse results and study the more general precoding strategy.

The rest of this paper is organized as follows: In Section II, we give a formal description of the system model. In Sections III and IV, we state our main results for the beamforming and precoding strategies, respectively.

*Notation:* For real-valued functions  $f(x)$ ,  $g(x)$ , let  $f(x) \in \omega(g(x))$  if (for all sufficiently large  $x$ )  $f(x) \geq kg(x)$ ,  $\forall k > 0$ ;  $f(x) \in O(g(x))$  if  $\exists a > 0$ ,  $f(x) \leq ag(x)$ ;  $f(x) \in o(g(x))$  if  $f(x) \leq eg(x)$ ,  $\forall \epsilon > 0$ ; and  $f(x) \sim g(x)$  if  $\lim_{x \rightarrow \infty} \frac{f(x)}{g(x)} = 1$ .  $\mathbf{A}^T$ ,  $\mathbf{A}^\dagger$  denote the transpose and the Hermitian transpose of matrix  $\mathbf{A}$ , respectively.  $\mathbf{A}^* = (\mathbf{A}^T)^\dagger$ .  $\mathbf{h} \simeq \mathcal{CN}(\mathbf{K})$  means that  $\mathbf{h}$  is a circulary-symmetric complex Gaussian random vector with covariance matrix  $\mathbf{K}$ .  $\mathbf{I}$  is the identity matrix, and  $\mathbf{0}$  is the all-zero matrix.  $|\mathcal{A}|$  is the cardinality of  $\mathcal{A}$ .

## II. PRELIMINARIES

### A. System Model

We consider a  $t \times 1$  MISO system. Denote the channel from transmitter antenna  $i$  to the receiver antenna by  $h_i$ , and let

$\mathbf{h} = [h_1 \dots h_t]^T \in \mathbb{C}^{t \times 1}$  represent the entire channel state. We assume that  $\mathbf{h} \simeq \text{CN}(\mathbf{I})$ .

We assume that the transmission symbol  $s$  is a discrete random variable with a uniform distribution on the set  $\{+1, -1\}$ .<sup>1</sup> For a fixed  $\mathbf{h}$ , we first consider the transmission of  $s$  via a beamforming vector  $\mathbf{x} \in \chi$ , where  $\chi = \{\mathbf{x} : \mathbf{x} \in \mathbb{C}^{t \times 1}, \|\mathbf{x}\| = 1\}$  is the set of all feasible beamforming vectors (We shall discuss the precoded transmission later on.). The channel input-output relationship with such a transmission strategy can be expressed as  $y = s\langle \mathbf{x}, \mathbf{h} \rangle \sqrt{P} + n$ , where  $y$  is the received signal, and the noise term  $n \simeq \text{CN}(1)$  is independent of  $\mathbf{h}$ . The corresponding signal-to-noise ratio (SNR) can be expressed as  $|\langle \mathbf{x}, \mathbf{h} \rangle|^2 P$ . Given  $\mathbf{h}$ , the conditional SER with a maximum-likelihood decoder is then  $Q(\sqrt{2}|\langle \mathbf{x}, \mathbf{h} \rangle|^2 P)$ , where  $Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^\infty e^{-\frac{u^2}{2}} du$ ,  $x \in \mathbb{R}$  is the Gaussian tail function.

When  $\mathbf{h}$  is random, we can choose a different beamforming vector for different  $\mathbf{h}$ . In this case, we are interested in the SER averaged over all possible channel states. Formally, consider an arbitrary (measurable) mapping  $\mathbf{m} : \mathbb{C}^t \rightarrow \chi$ . Then, the (average) SER with mapping  $\mathbf{m}$  can be expressed as

$$\text{SER}(\mathbf{m}) \triangleq \mathbb{E}[Q(\sqrt{2}|\langle \mathbf{m}(\mathbf{h}), \mathbf{h} \rangle|^2 P)]. \quad (1)$$

Let  $d(\mathbf{m}) = -\lim_{P \rightarrow \infty} [\text{SER}(\mathbf{m}) / \log P]$  as the diversity gain with  $\mathbf{m}$ , and  $g(\mathbf{m}) = [\lim_{P \rightarrow \infty} (\text{SER}(\mathbf{m}) P^{d(\mathbf{m})})]^{-1}$  as the array gain with  $\mathbf{m}$ , provided that both limits exist. The asymptotic  $P \rightarrow \infty$  performance of  $\mathbf{m}$  is then  $\text{SER}(\mathbf{m}) \sim [g(\mathbf{m}) P^{d(\mathbf{m})}]^{-1}$ .

As an extreme case, the transmitter may know  $\mathbf{h}$  perfectly, in which case we have a “full-CSIT system.” In such a scenario, we can choose an optimal beamforming vector, say  $\text{full}(\mathbf{h})$  for a given  $\mathbf{h}$ . We have  $|\langle \text{full}(\mathbf{h}), \mathbf{h} \rangle| \leq \|\mathbf{h}\|$ , and the upper bound is achievable by choosing  $\text{full}(\mathbf{h}) = \frac{\mathbf{h}}{\|\mathbf{h}\|}$ . This gives us  $\text{SER}(\text{full}) = \mathbb{E}[Q(\sqrt{2}\|\mathbf{h}\|^2 P)]$  with  $d(\text{full}) = t$ .

We now investigate the case where the transmitter has partial CSI via feedback from the receiver. Such a system can be modeled by a channel quantizer as we explain in the following.

### B. The channel quantizer

Let  $\mathcal{I} \in \{\{0\}, \{0, 1\}, \{0, 1, 2\}, \dots, \mathbb{N}\}$  be a possibly infinite index set. We use the notations  $\{a_n\}_{\mathcal{I}}$  and  $\{a_n : n \in \mathcal{I}\}$  interchangeably to represent a set whose elements are the real numbers  $a_n$ ,  $n \in \mathcal{I}$ . A similar definition holds for sets of vectors, collection of sets, etc.

Given  $\mathcal{I}$ , let  $\{\mathbf{x}_n\}_{\mathcal{I}}$  be a set of quantized beamforming vectors with  $\{\mathbf{x}_n\}_{\mathcal{I}} \subset \chi$ . Also, let  $\{\mathcal{E}_n\}_{\mathcal{I}}$  with  $\mathcal{E}_n \subset \mathbb{C}^t$ ,  $\forall n \in \mathbb{N}$  be a collection of mutually disjoint measurable subsets of  $\mathbb{C}^t$  with  $\bigcup_{n \in \mathcal{I}} \mathcal{E}_n = \mathbb{C}^t$ . Finally, let  $\{\mathbf{b}_n\}_{\mathcal{I}}$  be a collection of feedback binary codewords with  $\{\mathbf{b}_n\}_{\mathcal{I}} \subset \{0, 1\}^*$ , where  $\{0, 1\}^+ \triangleq \{0, 1, 00, 01, \dots\}$  is the set of all non-empty binary codewords. We assume that the code  $\{\mathbf{b}_n\}_{\mathcal{I}}$  is prefix-free, which implies in particular that  $\mathbf{b}_m \neq \mathbf{b}_n$  whenever  $m \neq n$ . We call the collection of triples  $\mathbf{q} \triangleq \{\mathbf{x}_n, \mathcal{E}_n, \mathbf{b}_n\}_{\mathcal{I}}$  a quantizer  $\mathbf{q}$  for the beamforming strategy.

<sup>1</sup>Our results can be extended to any finite constellation. We omit such a generalization here so as to highlight our quantization methods without dealing with the unnecessary technicalities of an arbitrary constellation.

This definition immediately induces a feedback transmission scheme that operates in the following manner: For a fixed channel state  $\mathbf{h}$ , the receiver feeds back the binary codeword  $\mathbf{b}_n$ , where the index  $n$  here satisfies  $\mathbf{h} \in \mathcal{E}_n$ . Such an index  $n$  always exists and is unique as  $\mathcal{E}_n$ ,  $n \in \mathbb{N}$  is a disjoint covering of  $\mathbb{C}^t$ . The transmitter recovers the index  $n$  and uses the corresponding beamforming vector  $\mathbf{x}_n$ . The recovery of  $n$  by the transmitter is always possible since  $\mathbf{b}_n$ s are distinct. We write  $\mathbf{q}(\mathbf{h}) = \mathbf{x}_n$  whenever  $\mathbf{h} \in \mathcal{E}_n$  to emphasize the quantization operation. We call the set  $\{\mathbf{x}_n\}_{\mathcal{I}}$  the quantizer (or beamforming) codebook.

For any  $\mathbf{b} \in \{0, 1\}^+$ , let  $L(\mathbf{b})$  denote the “length” of  $\mathbf{b}$ . For example,  $L(1) = 1$ ,  $L(01) = 2$ . A quantizer  $\mathbf{q}$  is called an FLQ if  $L(\mathbf{b}_m) = L(\mathbf{b}_n)$ ,  $\forall m, n \in \mathcal{I}$ . Otherwise, we call  $\mathbf{q}$  a VLQ. In either case, the rate of  $\mathbf{q}$  is  $R(\mathbf{q}) \triangleq \sum_{n \in \mathcal{I}} P(\mathbf{h} \in \mathcal{E}_n) L(\mathbf{b}_n)$ .

A quantizer  $\mathbf{q}$  is thus a mapping  $\mathbb{C}^t \rightarrow \{\mathbf{x}_n\}_{\mathcal{I}}$  supplied with a feedback binary codeword  $\mathbf{b}_n$  for each  $\mathbf{x}_n$ . Treated solely as a mapping, it is a special case of the mapping  $\mathbf{m} : \mathbb{C}^t \rightarrow \chi$  discussed in Section II-A with the requirement of a countable range  $\{\mathbf{x}_n\}_{\mathcal{I}}$ . We can therefore calculate the SER with  $\mathbf{q}$  as  $\text{SER}(\mathbf{q}) = \mathbb{E}[Q(\sqrt{2}|\langle \mathbf{q}(\mathbf{h}), \mathbf{h} \rangle|^2 P)]$ .

Our goal in this paper is to design low-rate VLQs that can achieve the full-CSIT diversity and array gains. Before we discuss our VLQ designs, let us reemphasize that we wish to achieve  $d(\text{full})$  and  $g(\text{full})$ , not  $\text{SER}(\text{full})$  (although there is no practically-significant difference between the two goals at high  $P$ ). The motivation behind this choice is not the hope that the diversity and array gains would be easier to work with. It is rather, as we shall prove in the following, the impossibility of achieving the full-CSIT SER with *any* quantizer. Note that one may achieve  $d(\text{full})$  and  $g(\text{full})$  while not achieving  $\text{SER}(\text{full})$  at any  $P$ . For example, suppose that a hypothetical quantizer  $\mathbf{q}'$  achieves  $\text{SER}(\mathbf{q}') = \text{SER}(\text{full}) + 1/P^{t+1}$ . Obviously we have  $\text{SER}(\mathbf{q}') > \text{SER}(\text{full})$ ,  $\forall P$ , while  $d(\mathbf{q}') = d(\text{full})$  and  $g(\mathbf{q}') = g(\text{full})$ .

### C. The impossibility of achieving $\text{SER}(\text{full})$

Let us first define the distortion function  $d(\mathbf{x}, \mathbf{h}) = Q(\sqrt{2}|\langle \mathbf{x}, \mathbf{h} \rangle|^2 P) - Q(\sqrt{2}\|\mathbf{h}\|^2 P)$ . For any given quantizer  $\mathbf{q}$ , we have  $\text{SER}(\mathbf{q}) = \text{SER}(\text{full}) + \mathbb{E}[d(\mathbf{q}(\mathbf{h}), \mathbf{h})]$ , and therefore, minimizing the expected distortion with  $\mathbf{q}$  is equivalent to minimizing the SER with  $\mathbf{q}$ . The following result is then merely the consequence of the countable nature of the quantizer structure and the properties of the distortion function  $d(\mathbf{x}, \mathbf{h})$  that is associated with the SER.

**Theorem 1.** *For any quantizer  $\mathbf{q}$ ,  $\text{SER}(\mathbf{q}) > \text{SER}(\text{full})$ ,  $\forall P > 0$ .*

*Proof:* Let  $\mathbf{q} = \{\mathbf{x}_n, \mathcal{E}_n, \mathbf{b}_n\}_{\mathcal{I}}$ . Since  $\sum_{n \in \mathcal{I}} P(\mathbf{h} \in \mathcal{E}_n) = 1$ ,  $\exists i \in \mathcal{I}$  such that  $P(\mathbf{h} \in \mathcal{E}_i) > 0$ . We then have

$$\begin{aligned} \text{SER}(\mathbf{q}) &= \text{SER}(\text{full}) + \sum_{n \in \mathcal{I}} \int_{\mathcal{E}_n} d(\mathbf{x}_n, \mathbf{h}) f(\mathbf{h}) d\mathbf{h} \\ &\geq \text{SER}(\text{full}) + \int_{\mathcal{E}_i} d(\mathbf{x}_i, \mathbf{h}) f(\mathbf{h}) d\mathbf{h}. \end{aligned}$$

For any  $P > 0$ , the distortion function  $d(\mathbf{x}_i, \mathbf{h})$  is positive almost everywhere (it is non-negative everywhere and it is zero only when  $\mathbf{h}$  is a member of the set  $\{\mathbf{h} : |\langle \mathbf{x}_i, \mathbf{h} \rangle|^2 = \|\mathbf{h}\|^2\}$ , which has probability measure zero). The integral of

an almost-everywhere-positive function on a set of positive measure is positive. Hence,  $\text{SER}(\mathbf{q}) > \text{SER}(\text{full}), \forall P > 0$ . ■

**Remark 1** (Comparison with outage probability [22]). For the same  $t \times 1$  MISO system considered in this paper, and for a given target data transmission rate, let  $\text{OUT}(\mathbf{q})$  and  $\text{OUT}(\text{full})$  denote the outage probability with a (beamforming) quantizer  $\mathbf{q}$  and the full-CSIT outage probability, respectively. Then, similarly, there is a distortion function  $D(\mathbf{x}, \mathbf{h})$  that satisfies  $\text{OUT}(\mathbf{q}) = \text{OUT}(\text{full}) + E[D(\mathbf{x}, \mathbf{h})]$ . However, unlike the distortion function  $d(\mathbf{x}, \mathbf{h})$  for the SER, for any  $\mathbf{x} \in \mathcal{X}$ , the function  $D(\mathbf{x}, \mathbf{h})$  is zero on a set of positive measure [22]. Hence, the natural analogue of Theorem 1, i.e. the claim that  $\text{OUT}(\mathbf{q}) > \text{OUT}(\text{full})$  for any quantizer  $\mathbf{q}$ , may not hold in the case of outage probability. In fact, it is possible to design finite-rate VLQs that can achieve  $\text{OUT}(\text{full})$  at any  $P$  [22]. Hence, at least in the context of limited feedback, the SER and outage probability performance measures exhibit fundamentally different behaviors. ■

According to Theorem 1, we have no hope in achieving  $\text{SER}(\text{full})$ . The next most important question is how to design quantizers that can achieve  $d(\text{full})$  and  $g(\text{full})$  if possible. In this context, it is well-known that finite-rate FLQs cannot achieve these full-CSIT diversity and array gains [6]. In this paper, we design VLQs that can achieve these gains with a feedback rate of 1 bit per channel state asymptotically as  $P \rightarrow \infty$ . This is a significant improvement over FLQs that require infinite rate to achieve the same performance. We first discuss how to design such quantizers for the beamforming strategy; the precoding case will be discussed afterwards.

### III. VLQS FOR BEAMFORMING

We start with the design of the quantizer encoding regions for a given beamforming codebook. We first recall the standard encoding rule for FLQs and discuss why it will not work in the case of VLQs. We then modify the standard encoder to come up with a new encoder that will allow us to design good VLQs.

#### A. Encoding

Let  $\mathcal{B} = \{\mathbf{x}_n\}_{\mathcal{I}}$  be a finite-cardinality beamforming codebook. A standard practice (see e.g. [6], [12]) is to work with the quantizer  $\mathbf{q}_{\mathcal{B}}(\mathbf{h}) \triangleq \arg \max_{\mathbf{x} \in \mathcal{B}} |\langle \mathbf{x}, \mathbf{h} \rangle|$ , which chooses the beamforming vector (with ties broken arbitrarily) in  $\mathcal{B}$  that is “closest” to  $\mathbf{h}$ . One way to design a VLQ might be to keep this standard encoding rule for FLQs but use a variable-length code instead of a fixed-length code. On the other hand, the standard encoding rule results in quantization cells with roughly equal probability  $1/|\mathcal{B}|$  when  $|\mathcal{B}|$  is large. In such a scenario, the rate of an optimal variable-length code will roughly be the same as the rate  $\lceil \log_2 |\mathcal{B}| \rceil$  of the fixed-length code [22]. We thus consider an alternate encoding strategy that can benefit from variable-length codes.

The standard encoder always picks the best beamforming vector in  $\mathcal{B} = \{\mathbf{x}_n\}_{\mathcal{I}}$  that is closest to  $\mathbf{h}$ . We do not have to be this precise if our goal is to achieve the diversity and array gains provided by  $\mathcal{B}$ . For example, we do not need to

distinguish between two beamforming vectors given that both provide an SER of at most  $o(1/P^t)$ ; preferring one vector over the other will not affect the diversity and array gains of the system as the best possible decay of the SER is  $O(1/P^t)$ .

With this observation, for a given beamforming codebook  $\mathcal{B}$ , we consider a variable-length quantizer  $\mathbf{q}_{\mathcal{B}}^v$  that operates as follows. Let  $\beta = (t+1) \log P$ .

- If  $|\langle \mathbf{x}_i, \mathbf{h} \rangle|^2 P \geq \beta, \forall i \in \mathcal{I}$ , then  $\mathbf{q}_{\mathcal{B}}^v(\mathbf{h})$  feeds back the binary codeword 0, and we set  $\mathbf{q}_{\mathcal{B}}^v(\mathbf{h}) = \mathbf{x}_0$ .
- Otherwise, if  $\exists i \in \mathcal{I}$  with  $|\langle \mathbf{x}_i, \mathbf{h} \rangle|^2 P < \beta$ , then  $\mathbf{q}_{\mathcal{B}}^v$  feeds back the concatenation of the binary codeword 1 and the binary codeword of length  $\lceil \log_2 |\mathcal{B}| \rceil$  bits that represents the index, say  $j \in \mathcal{I}$ , of the beamforming vector that results in the maximum SNR. We set  $\mathbf{q}_{\mathcal{B}}^v(\mathbf{h}) = \mathbf{q}_{\mathcal{B}}(\mathbf{h})$ .

Let us now analyze the performance of  $\mathbf{q}_{\mathcal{B}}^v$ .

**Proposition 1.** *For any finite-cardinality beamforming codebook  $\mathcal{B}$ , we have*

$$\text{SER}(\mathbf{q}_{\mathcal{B}}^v) \leq \text{SER}(\mathbf{q}_{\mathcal{B}}) + \frac{1}{P^{t+1}}, \quad (2)$$

$$R(\mathbf{q}_{\mathcal{B}}^v) \leq 1 + \frac{(t+1)|\mathcal{B}| \log_2(4|\mathcal{B}|) \log P}{P}. \quad (3)$$

*Proof:* We first prove the upper bound on  $\text{SER}(\mathbf{q}_{\mathcal{B}}^v)$ . Let  $d(\mathbf{h}) = Q(\sqrt{2}|\langle \mathbf{q}_{\mathcal{B}}^v(\mathbf{h}), \mathbf{h} \rangle|^2 P) - Q(\sqrt{2}|\langle \mathbf{q}_{\mathcal{B}}(\mathbf{h}), \mathbf{h} \rangle|^2 P)$ ,  $\mathcal{E} = \{\mathbf{h} : \exists i \in \mathcal{I}, |\langle \mathbf{x}_i, \mathbf{h} \rangle|^2 P < \beta\}$ , and  $\mathcal{E}^c = \mathbb{C}^t - \mathcal{E}$ . By the definition of  $\mathbf{q}_{\mathcal{B}}^v$ , we have  $E[d(\mathbf{h})|\mathbf{h} \in \mathcal{E}] = 0$ . Also,  $E[d(\mathbf{h})|\mathbf{h} \in \mathcal{E}^c] \leq E[Q(\sqrt{2}|\langle \mathbf{q}_{\mathcal{B}}^v(\mathbf{h}), \mathbf{h} \rangle|^2 P)] \leq E[Q(\sqrt{2(t+1) \log P})] \leq P^{-(t+1)}$ , where the second inequality follows since  $|\langle \mathbf{q}_{\mathcal{B}}^v(\mathbf{h}), \mathbf{h} \rangle|^2 P \geq (t+1) \log P, \forall \mathbf{h} \in \mathcal{E}^c$ , and the last inequality follows from the bound  $Q(x) \leq e^{-x^2}, x \geq 0$ . Combining the two conditional expectations of  $d(\mathbf{h})$ , we obtain  $E[d(\mathbf{h})] \leq P^{-(t+1)}$ . Substituting this to the obvious identity  $\text{SER}(\mathbf{q}_{\mathcal{B}}^v) = \text{SER}(\mathbf{q}_{\mathcal{B}}) + E[d(\mathbf{h})]$ , we obtain (2).

We now prove (3). We have

$$R(\mathbf{q}_{\mathcal{B}}^v) = P(\mathbf{h} \in \mathcal{E}^c) + (1 + \lceil \log_2 |\mathcal{B}| \rceil)P(\mathbf{h} \in \mathcal{E}) \quad (4)$$

$$\leq 1 + \log_2(4|\mathcal{B}|)P(\mathbf{h} \in \mathcal{E}) \quad (5)$$

$$\leq 1 + \log_2(4|\mathcal{B}|) \sum_{i \in \mathcal{I}} P(|\langle \mathbf{x}_i, \mathbf{h} \rangle|^2 P \leq \beta) \quad (6)$$

$$= 1 + \log_2(4|\mathcal{B}|)|\mathcal{B}|[1 - \exp(-\frac{\beta}{P})] \quad (7)$$

$$\leq 1 + \log_2(4|\mathcal{B}|)|\mathcal{B}| \frac{\beta}{P}, \quad (8)$$

where the first inequality follows since  $P(\mathbf{h} \in \mathcal{E}^c) \leq 1$  and  $1 + \lceil \log_2 |\mathcal{B}| \rceil \leq 1 + \log_2 |\mathcal{B}| + 1 = \log_2(4|\mathcal{B}|)$ . The second inequality follows from a union bound. This proves (3). ■

Note that the maximum diversity gain with any quantizer is  $t$ . Hence,  $\text{SER}(\mathbf{q}_{\mathcal{B}}) \simeq g(\mathbf{q}_{\mathcal{B}})P^{-d}$  for some  $d \leq t$ . Since the second term in the upper bound in (2) decays faster than  $\frac{1}{P^t}$ , the diversity and array gains with  $\mathbf{q}_{\mathcal{B}}^v$  is the same with those of  $\mathbf{q}_{\mathcal{B}}$ . Moreover, according to (3), we have  $R(\mathbf{q}_{\mathcal{B}}^v) \rightarrow 1$  as  $P \rightarrow \infty$ . This is a significant improvement over a rate- $\lceil \log_2 |\mathcal{B}| \rceil$  FLQ for codebook  $\mathcal{B}$ , especially when  $|\mathcal{B}|$  is large.

We now claim that, for any arbitrary function  $f(P) \in \omega(1)$ , there is a VLQ that can achieve  $d(\text{full})$  and  $g(\text{full})$  with rate  $1 + f(P) \frac{\log P}{P}$ . We provide here an outline of the strategy to prove this result. Motivated by (3), we consider  $P$ -dependent

codebooks  $\mathcal{B}_P$  that satisfy  $1 + (t+1)|\mathcal{B}_P|\lceil \log_2 |\mathcal{B}_P| \rceil = f(P)$ . In such a scenario,  $|\mathcal{B}_P| \in \omega(1)$ , or equivalently, we use codebooks with larger and larger cardinality as  $P \rightarrow \infty$ . If we can design the codebooks  $\mathcal{B}_P$ ,  $P > 0$  well enough, the quantizer  $q_{\mathcal{B}_P}^v$  will then achieve  $d(\text{full})$  and  $g(\text{full})$  with rate  $1 + f(P)\frac{\log P}{P}$ , as claimed. Obviously, for this strategy to work, we need “good” codebook designs. We show the existence of such codebooks in the following.

### B. Existence of good codebooks

The following result is a restatement of [22, Proposition 2].

**Proposition 2.** *For every sufficiently small  $\delta > 0$ , there is a codebook  $\mathcal{B}_\delta$  with*

$$|\mathcal{B}_\delta| \leq C_0 \delta^{-2t}, \quad (9)$$

and

$$\sqrt{\mathbf{h}} \in \chi, \exists \mathbf{x} \in \mathcal{B}_\delta, |\langle \mathbf{x}, \sqrt{\mathbf{h}} \rangle|^2 \geq 1 - \delta, \quad (10)$$

where  $C_0$  is a  $\delta$ -independent constant.

**Remark 2.** This existence result is good enough for our purposes. Still, let us note that an explicit construction for  $\mathcal{B}_\delta$  is also available for implementation purposes. We refer the interested reader to [22, Section III.B]. ■

In the following proposition, we calculate the SER with  $\mathcal{B}_\delta$ .

**Proposition 3.** *For every sufficiently small  $\delta > 0$ , we have*

$$\text{SER}(q_{\mathcal{B}_\delta}) \leq \text{SER}(\text{full})(1 + 2t\delta). \quad (11)$$

*Proof:* According to Proposition (2), the SNR provided by  $\mathcal{B}_\delta$  is at least  $\|\mathbf{h}\|^2(1 - \delta)P$  for any given channel state  $\mathbf{h}$ . As a result,  $\text{SER}(q_{\mathcal{B}_\delta}) \leq \int_0^\infty Q(\sqrt{2x(1 - \delta)P})x^{t-1}e^{-x}/\Gamma(t)dx$ , where the integration variable  $x$  corresponds to a realization of the random variable  $\|\mathbf{h}\|^2$ . With a change of variables  $u = x(1 - \delta)$ , we obtain  $\text{SER}(q_{\mathcal{B}_\delta}) \leq (1 - \delta)^{-t} \int_0^\infty Q(\sqrt{2uP})u^{t-1}e^{-u/(1-\delta)}/\Gamma(t)du$ . For the factor  $(1 - \delta)^{-t}$  of the integral, we have  $(1 - \delta)^{-t} = 1 + t\delta + O(\delta^2) \leq 1 + 2t\delta$  for any sufficiently small  $\delta$ . For the integral itself, we use the bound  $e^{-\frac{u}{1-\delta}} \leq e^{-u}$ , which makes the integral equal to  $\text{SER}(\text{full})$ . This concludes the proof. ■

Hence, for sufficiently small  $\delta$ , the codebook  $\mathcal{B}_\delta$  can provide the full-diversity gain. Making  $\delta$  even smaller, it can also provide an array gain that is arbitrarily close to  $g(\text{full})$ .

### C. The main achievability result

We can now proceed with the strategy outlined at the end of Section III-A. The following is the main result of this section.

**Theorem 2.** *For any function  $f(P) \in \omega(1)$ , there is a quantizer  $q$  with  $d(q) = d(\text{full})$ ,  $g(q) = g(\text{full})$ , and  $R(q) \leq 1 + f(P)\frac{\log P}{P}$  for all sufficiently large  $P$ .*

*Proof:* By (3) and (9), for every sufficiently small  $\delta$ , we have  $R(q_{\mathcal{B}_\delta}) \leq 1 + \phi(\delta)\frac{\log P}{P}$ , where  $\phi(\delta) \triangleq (t+1)C_0/\delta^{2t} \times \log_2(4C_0/\delta^{2t})$ . Hence, for given  $f(P) \in \omega(1)$ , we choose  $\delta(P) \triangleq \phi^{-1}(f(P))$ . Since  $f(P) \rightarrow \infty$ ,  $\delta(P)$  is  $o(1)$ , well-defined, and positive for all sufficiently large  $P$ . The quantizer

$q = q_{\mathcal{B}_{\delta(P)}}^v$  now satisfies the theorem’s statement. Indeed, by (2) and (11), we have  $\text{SER}(q) \leq \text{SER}(\text{full})(1 + 2t\delta(P)) + 1/P^{t+1}$ . Since  $\delta(P) \in o(1)$ , we have  $d(q) = d(\text{full})$ ,  $g(q) = g(\text{full})$ . The upper bound on  $R(q)$  is obvious. ■

In particular, for any  $f(P) \in o(\frac{P}{\log P})$  (e.g.  $f(P) = \log P$ ), the full-CSIT diversity and array gains can be achieved with a feedback rate of 1 bit per channel state asymptotically. The question is now to determine the minimum rate that guarantees the full-CSIT gains. We discuss this problem next.

### D. Necessary conditions for achieving $d(\text{full})$ and $g(\text{full})$

It is difficult to determine the exact asymptotic rate that guarantees the full-CSIT gains. Instead, we provide bounds. Note that by Theorem 2, a quantization rate of  $1 + \frac{\omega(1)\log P}{P}$  is sufficient for the full-CSIT gains. We prove a partial converse by showing that a quantization rate of  $1 + \frac{t\log P}{13P}$  is necessary.

**Theorem 3.** *For any quantizer  $q$ , if  $d(q) = d(\text{full})$ ,  $g(q) = g(\text{full})$ , then  $R(q) \geq 1 + \frac{t\log P}{13P}$  for all sufficiently large  $P$ .*

*Proof:* Let us first (roughly) identify the class of quantizers that can achieve  $d(\text{full})$  and  $g(\text{full})$ . Let  $q = \{\mathbf{x}_n, \mathcal{E}_n, \mathbf{b}_n\}_{\mathcal{I}}$ . The case  $|\mathcal{I}| = 1$  corresponds to open-loop system with no CSIT, in which case we have  $d(q) \leq 1 < d(\text{full})$ . The case  $|\mathcal{I}| = 2$  corresponds to a quantizer whose codebook consists of 2 beamforming vectors. In such a scenario,  $d(q) = 2$  is achievable with  $R(q) = 1$ . On the other hand, for  $t = 2$ , either  $d(q) < d(\text{full}) = 2$ , or if  $d(q) = d(\text{full}) = 2$ , then  $g(q) < g(\text{full})$  as one cannot achieve the full-CSIT array gain with only 2 beamforming vectors. For  $t \geq 3$ , we simply have  $d(q) \leq 2 < 3 \leq d(\text{full})$ .

It is thus sufficient to consider the case  $|\mathcal{I}| \geq 3$  for the class of quantizers that can achieve  $d(\text{full})$  and  $g(\text{full})$ . Hence, let  $q = \{\mathbf{x}_n, \mathcal{E}_n, \mathbf{b}_n\}_{\mathcal{I}}$  with  $|\mathcal{I}| \geq 3$  and suppose that for every  $P_0 \in \mathbb{R}$ , there exists  $P \geq P_0$  such that  $R(q) < 1 + \frac{t\log P}{13P}$ . We will conclude the proof of the theorem by showing that the strict inequality  $d(q) < d(\text{full})$  then holds.

Let  $R = \frac{t\log P}{13P}$ . Since,  $|\mathcal{I}| \geq 3$ , there is an index  $i \in \mathcal{I}$  with  $L(\mathbf{b}_i) \geq 1$  and  $L(\mathbf{b}_j) \geq 2$ ,  $\forall j \in \mathcal{I} - \{i\}$  (Otherwise, there would be two binary codewords of length 1, which is impossible for a prefix-free code that contains at least 3 codewords.). Then, we have  $P(\mathbf{h} \in \mathcal{E}_i) \geq 1 - R$  (Otherwise, the chain of inequalities

$$1 + R > R(q) \quad (12)$$

$$= \sum_{n \in \mathcal{I}} P(\mathbf{h} \in \mathcal{E}_n)L(\mathbf{b}_n) \quad (13)$$

$$= P(\mathbf{h} \in \mathcal{E}_i)L(\mathbf{b}_i) + \sum_{j \in \mathcal{I} - \{i\}} P(\mathbf{h} \in \mathcal{E}_j)L(\mathbf{b}_j) \quad (14)$$

$$\geq P(\mathbf{h} \in \mathcal{E}_i) + 2 \sum_{j \in \mathcal{I} - \{i\}} P(\mathbf{h} \in \mathcal{E}_j) \quad (15)$$

$$= P(\mathbf{h} \in \mathcal{E}_i) + 2(1 - P(\mathbf{h} \in \mathcal{E}_i)) \quad (16)$$

$$= 2 - P(\mathbf{h} \in \mathcal{E}_i) \quad (17)$$

$$\geq 1 + R \quad (18)$$

leads to a contradiction.). Without loss of generality, suppose that  $P(\mathbf{h} \in \mathcal{E}_0) \geq 1 - R$ . Then, with  $f(\mathbf{h})$  representing the probability density function of  $\mathbf{h}$ , we have

$$\text{SER}(q) = \sum_{n \in \mathcal{I}} \int_{\mathcal{E}_n} Q(\sqrt{2|\langle \mathbf{x}_n, \mathbf{h} \rangle|^2 P})f(\mathbf{h})d\mathbf{h} \quad (19)$$

$$\geq \int_{\mathcal{E}_0} Q(\sqrt{2|\langle \mathbf{x}_0, \mathbf{h} \rangle|^2 P}) f(\mathbf{h}) d\mathbf{h} \quad (20)$$

According to [25, Theorem 2.1], there is a constant  $C_1 > 0$  such that  $Q(x) \geq C_1 \exp(-x^2)$ ,  $\forall x \in \mathbb{R}$ . Substituting this lower bound, we obtain

$$\text{SER}(\mathbf{q}) \geq g(\mathcal{E}_0, \mathbf{x}_0), \quad (21)$$

where

$$g(\mathcal{E}, \mathbf{x}) = C_1 \int_{\mathcal{E}} \exp(-2|\langle \mathbf{x}, \mathbf{h} \rangle|^2 P) f(\mathbf{h}) d\mathbf{h} \quad (22)$$

Now, let  $\mathfrak{E} = \{\mathcal{E} \subset \mathbb{C}^t : P(\mathbf{h} \in \mathcal{E}) \geq 1 - R\}$ . Clearly,  $\mathcal{E}_0 \in \mathfrak{E}$ , and therefore

$$\text{SER}(\mathbf{q}) \geq \inf_{\mathcal{E} \in \mathfrak{E}} \inf_{\mathbf{x} \in \chi} g(\mathcal{E}, \mathbf{x}). \quad (23)$$

Note that for any  $t \times t$  unitary matrix  $\mathbf{U}$ , we have  $g(\mathcal{E}, \mathbf{x}) = g(\mathbf{U}\mathcal{E}, \mathbf{U}\mathbf{x})$ ,  $\forall \mathcal{E} \in \mathfrak{E}$ ,  $\forall \mathbf{x} \in \chi$ , where  $\mathbf{U}\mathcal{E} = \{\mathbf{Uh} : \mathbf{h} \in \mathcal{E}\}$  denotes the translate of the set  $\mathcal{E}$  by  $\mathbf{U}$ . With this property in mind, we consider a fixed vector  $\mathbf{y} \in \chi$ . For a given  $\mathbf{x} \in \chi$ , let the unitary matrix  $\mathbf{U}_{\mathbf{x}}$  satisfy  $\mathbf{x} = \mathbf{U}_{\mathbf{x}}\mathbf{y}$ . We have  $g(\mathcal{E}, \mathbf{x}) = g(\mathbf{U}_{\mathbf{x}}^{\dagger}\mathcal{E}, \mathbf{U}_{\mathbf{x}}^{\dagger}\mathbf{x}) = g(\mathbf{U}_{\mathbf{x}}^{\dagger}\mathcal{E}, \mathbf{y})$ . Since  $\mathbf{U}_{\mathbf{x}}^{\dagger}\mathcal{E} \in \mathfrak{E}$ , we have  $g(\mathcal{E}, \mathbf{x}) \geq \inf_{\mathcal{E}' \in \mathfrak{E}} g(\mathcal{E}', \mathbf{y})$ . Since this inequality holds for arbitrary  $\mathbf{x}$  and  $\mathcal{E}$ , we obtain  $\inf_{\mathcal{E} \in \mathfrak{E}} \inf_{\mathbf{x} \in \chi} g(\mathcal{E}, \mathbf{x}) \geq \inf_{\mathcal{E} \in \mathfrak{E}} g(\mathcal{E}, \mathbf{y})$ . Choosing e.g.  $\mathbf{y} = [1 0 \dots 0]^T$  then gives us

$$\text{SER}(\mathbf{q}) \geq C_1 \inf_{\mathcal{E} \in \mathfrak{E}} \int_{\mathcal{E}} \exp(-2|h_1|^2 P) f(\mathbf{h}) d\mathbf{h} \quad (24)$$

The expression in the lower bound is an optimization problem of the form ‘‘Minimize  $\int_{\mathcal{E}} \exp(-2|h_1|^2 P) d\mu$ , subject to  $\mu(\mathcal{E}) \geq 1 - R$ ,’’ where  $\mu$  is a probability measure. According to [26], there is a minimizer of the form  $\mathcal{E}' = \{\mathbf{h} : |h_1|^2 \geq r\}$ , where  $r$  is a positive real number with  $\mu(\mathcal{E}') = 1 - R$ . In other words, one forms the solution set  $\mathcal{E}'$  by starting with the points where the integrand  $\exp(-2|h_1|^2 P)$  takes its minimal values and then progressively adds more points until the measure of  $\mathcal{E}'$  is equal to  $r$ . We have

$$\int_{\{\mathbf{h}:|h_1|^2 \geq r\}} f(\mathbf{h}) d\mathbf{h} = 1 - R, \quad (25)$$

or, equivalently  $\int_r^\infty e^{-x} dx = 1 - R$ . Solving for  $r$ , we obtain  $r = -\log(1 - R)$ , and thus

$$\text{SER}(\mathbf{q}) \geq C_1 \int_{-\log(1-R)}^\infty \exp(-2xP) e^{-x} dx \quad (26)$$

$$= \frac{C_1 \exp[(1+2P)\log(1-R)]}{1+2P} \quad (27)$$

$$\geq \frac{C_1 \exp(-6PR)}{3P}, \quad (28)$$

where the last inequality follows since  $1 + 2P \leq 3P$ ,  $\forall P \geq 1$ , and  $-\log(1 - R) \leq 2R$  for sufficiently small  $R$ . Now, substituting  $R = \frac{t \log P}{13P}$ , the lower bound is proportional to  $P^{-1 - \frac{6t}{13}}$ . This means  $d(\mathbf{q}) \leq 1 + \frac{6t}{13} < t$ ,  $\forall t \geq 2$ , if  $d(\mathbf{q})$  ever exists. This concludes the proof. ■

We thus say that the necessary and sufficient feedback rate that guarantees the full-CSIT gains is  $1 + \omega(1) \frac{\log P}{P}$ , up to  $o(1)$  multipliers in the  $P$ -dependent term  $\omega(1) \frac{\log P}{P}$ . The fundamental question of whether the rate of  $1 + \omega(1) \frac{\log P}{P}$  is necessary will remain an open problem.

#### IV. VLQS FOR PRECODING

So far, we have designed VLQs that can achieve the full-CSIT gains with rate  $1 + \frac{\omega(1) \log P}{P}$ . The question is whether or not the decay rate of the term  $\frac{\omega(1) \log P}{P}$  can be improved using more sophisticated data transmission strategies. The answer is yes, and as a proof of concept, we consider linear precoding of a complex orthogonal space-time block code.

##### A. Linear precoding of complex orthogonal designs

Consider a fixed  $\mathbf{h}$ . We consider the transmission of  $k$  symbols  $s_1, \dots, s_k$  over  $n$  time slots via an  $n \times t$  complex orthogonal design  $\mathbf{S}$  (with  $\mathbf{S}^{\dagger}\mathbf{S} = \sum_k |s_k|^2 \mathbf{I}$ ,  $\forall s_1, \dots, s_k \in \mathbb{C}$ ) multiplied by a precoding matrix  $\mathbf{X} \in \mathcal{X}$ , where  $\mathcal{X} = \{\mathbf{X} \in \mathbb{C}^{t \times t} : \|\mathbf{X}\| \leq 1\}$ . The channel input-output relationship can be expressed as  $\mathbf{y} = \mathbf{SXh} \sqrt{P/r} + \mathbf{n}$ , where  $r = k/n$  is the space-time code rate,  $\mathbf{y}$  is the received signal vector, and  $\mathbf{n} \sim \mathcal{CN}(\mathbf{I})$ . Due to the orthogonality of the code, the per-symbol conditional SNR is  $\|\mathbf{Xh}\|^2 P/r$ . It follows that the conditional SER (for BPSK) is given by  $Q(\sqrt{2\|\mathbf{Xh}\|^2 P/r})$ . In particular, if, for example,  $\mathbf{X} = \mathbf{xx}^{\dagger}$  for some  $\mathbf{x} \in \chi$ , the conditional SER is  $Q(\sqrt{2|\langle \mathbf{x}, \mathbf{h} \rangle|^2 P/r})$ . Hence, up to the  $\frac{1}{r}$ -scaling of the SNR, using the precoding matrix  $\mathbf{xx}^{\dagger}$  is equivalent to beamforming along  $\mathbf{x}$ .<sup>2</sup>

When  $\mathbf{h}$  is random, we consider a mapping  $\mathbf{M} : \mathbb{C}^{t \times 1} \rightarrow \mathcal{X}$ , and let  $\text{SER}_r(\mathbf{M}) \triangleq E[Q(\sqrt{2\|\mathbf{M}(\mathbf{h})\mathbf{h}\|^2 P/r})]$  denote the SER with  $\mathbf{M}$ . The diversity and array gains  $d(\mathbf{M})$  and  $g(\mathbf{M})$  with mapping  $\mathbf{M}$  can be defined in the same manner as in Section II-A. The full-CSIT mapping  $\text{FULL}(\mathbf{h}) \triangleq \frac{\mathbf{hh}^{\dagger}}{\|\mathbf{h}\|^2}$  provides the minimum possible SER,  $\text{SER}_r(\text{FULL}) = E[Q(\sqrt{2\|\mathbf{h}\|^2 P/r})]$ , by beamforming along  $\frac{\mathbf{h}}{\|\mathbf{h}\|}$  for every  $\mathbf{h}$ .<sup>3</sup>

##### B. Main results

The quantizer definition in Section II-B for beamforming extends to precoding in an obvious manner. We let the collection of triples  $\mathbf{Q} \triangleq \{\mathbf{X}_n, \mathcal{E}_n, \mathbf{b}_n\}_{\mathcal{I}}$  be a channel quantizer for the precoding strategy, where  $\{\mathbf{X}_n\}_{\mathcal{I}} \subset \mathcal{X}$  with  $\{\mathcal{E}_n\}_{\mathcal{I}}$  and  $\{\mathbf{b}_n\}_{\mathcal{I}}$  satisfying the properties described in Section II-B.

It is straightforward to show that Theorem 1 holds for the precoding case as well. We thus construct quantized precoders that can achieve the full-CSIT diversity and array gains. We omit here a detailed exposition of the construction due to lack of space and since the ideas are very similar to the ones we have applied for beamforming.

We consider the precoding codebook  $\mathcal{C}_{\delta} \triangleq \{\mathbf{I}/\sqrt{t}\} \cup \{\mathbf{xx}^{\dagger} : \mathbf{x} \in \mathcal{B}_{\delta}\}$  that consists of the identity precoder and the precoder equivalents of the beamforming vectors in  $\mathcal{B}_{\delta}$ . We then construct a VLQ  $\mathbf{Q}_{\mathcal{C}_{\delta}}^v$  that operates as follows:

- If  $\|\mathbf{I}/\sqrt{t}\mathbf{h}\|^2 P \geq \delta^{-1}$ , then  $\mathbf{Q}_{\mathcal{C}_{\delta}}^v$  feeds back the binary codeword 0, and we set  $\mathbf{Q}_{\mathcal{C}_{\delta}}^v(\mathbf{h}) = \mathbf{I}/\sqrt{t}$ .

<sup>2</sup>In fact, using any precoding matrix of the form  $\mathbf{zx}^{\dagger}$ ,  $\mathbf{z} \in \chi$  is equivalent to beamforming along  $\mathbf{x}$ . For example,  $\mathbf{z} = [1 \dots 1]^T$  gives us the matrix  $[\mathbf{x} \dots \mathbf{x}]^{\dagger}$ . We have picked  $\mathbf{z} = \mathbf{x}$  (which yields  $\mathbf{xx}^{\dagger}$ ) for a simpler notation.

<sup>3</sup>For  $t = 2$  and the Alamouti code, we have  $r = 1$ , and thus  $\text{SER}_1(\text{FULL}) = \text{SER}(\text{full})$ . For  $t \geq 3$ , due to the non-existence of rate-1 complex-orthogonal designs, we have  $\text{SER}_r(\text{FULL}) < \text{SER}(\text{full})$ : A ‘‘part’’ of what is lost in terms of data rate translates to a constant power gain.

- Otherwise,  $\mathbf{Q}_{C_\delta}^v$  feeds back the concatenation of the binary codeword 1 and the binary codeword of length  $\lceil \log_2 |\mathcal{B}_\delta| \rceil$  bits that represents  $\mathbf{q}_{\mathcal{B}_\delta}(\mathbf{h})$ . We set  $\mathbf{Q}_{C_\delta}^v(\mathbf{h}) = [\mathbf{q}_{\mathcal{B}_\delta}(\mathbf{h})][\mathbf{q}_{\mathcal{B}_\delta}(\mathbf{h})]^\dagger$ .

In the following, we analyze the performance of  $\mathbf{Q}_{C_\delta}^v$ .

**Proposition 4.** *For every sufficiently small  $\delta$ ,*

$$\text{SER}_r(\mathbf{Q}_{C_\delta}^v) \leq \text{SER}_r(\text{FULL})(1 + 2t\delta) + \delta/P^t, \quad (29)$$

$$R(\mathbf{Q}_{C_\delta}^v) \leq 1 + C_2 \delta^{-t} \log(\delta^{-1})/P^t, \quad (30)$$

where  $C_2$  is a constant that is independent of  $\delta$  and  $P$ .

*Proof:* By construction, the precoding quantizer  $\mathbf{Q}_{C_\delta}^v$  provides the same SNR as the beamforming quantizer  $\mathbf{q}_{\mathcal{B}_\delta}$  except when  $\|\mathbf{h}\|^2 P \geq t\delta^{-1}$ , in which case it provides an SNR of  $\|\mathbf{h}\|^2 P/t$ . This gives us

$$\text{SER}_r(\mathbf{Q}_{C_\delta}^v) \leq \text{SER}_r(\mathbf{q}_{\mathcal{B}_\delta}) + \int_{\frac{t}{\delta P}}^{\infty} Q(\sqrt{2xP/t}) \frac{x^{t-1} e^{-x}}{\Gamma(t)} dx \quad (31)$$

$$\leq \text{SER}_r(\mathbf{q}_{\mathcal{B}_\delta}) + \int_{\frac{t}{\delta P}}^{\infty} \exp(-xP/t) \frac{x^{t-1}}{\Gamma(t)} dx \quad (32)$$

$$= \text{SER}_r(\mathbf{q}_{\mathcal{B}_\delta}) + \frac{t^t}{P^t} \int_{\delta^{-1}}^{\infty} \frac{u^{t-1} e^{-u}}{\Gamma(t)} du \quad (33)$$

$$= \text{SER}_r(\mathbf{q}_{\mathcal{B}_\delta}) + \frac{t^t}{P^t} e^{-\delta^{-1}} \sum_{k=0}^{t-1} \frac{\delta^{-k}}{k!} \quad (34)$$

$$\leq \text{SER}_r(\mathbf{q}_{\mathcal{B}_\delta}) + \frac{\delta}{P^t}, \quad (35)$$

where  $\text{SER}_r(\mathbf{q}_{\mathcal{B}_\delta}) \triangleq E[Q(\sqrt{2|\langle \mathbf{q}_{\mathcal{B}_\delta}(\mathbf{h}), \mathbf{h} \rangle|^2 P/r})]$ , and the last inequality holds for all sufficiently small  $\delta$ . Then, (29) follows since (using the same arguments as in Proposition 3)  $\text{SER}_r(\mathbf{q}_{\mathcal{B}_\delta}) \leq \text{SER}_r(\text{FULL})(1 + 2t\delta)$ .

For (30), note that  $\mathbf{Q}_{C_\delta}^v$  feeds back 1 bit if  $\|\mathbf{h}\|^2 P \geq t\delta^{-1}$ , and it feeds back  $1 + \lceil \log_2 |\mathcal{B}_\delta| \rceil$  bits if  $\|\mathbf{h}\|^2 P < t\delta^{-1}$ . The latter event has probability

$$\int_0^{\frac{t}{\delta P}} \frac{x^{t-1} e^{-x}}{\Gamma(t)} dx \leq \int_0^{\frac{t}{\delta P}} \frac{x^{t-1}}{\Gamma(t)} dx = \frac{t^t \delta^{-t}}{\Gamma(t+1) P^t}. \quad (36)$$

Combining this with the bound  $|\mathcal{B}_\delta| \leq C_0 \delta^{-2t}$  in (9), we obtain

$$R(\mathbf{Q}_{C_\delta}^v) \leq 1 + (1 + \lceil C_0 \delta^{-2t} \rceil) \frac{t^t \delta^{-t}}{\Gamma(t+1) P^t}, \quad (37)$$

After some straightforward manipulations, this yields the same upper bound as in the statement of the proposition. ■

This immediately leads to the following theorem. The proof is omitted since it is very similar to the proof of Theorem 2.

**Theorem 4.** *For any function  $f(P) \in \omega(1)$ , there is a precoding quantizer  $\mathbf{Q}$  with  $d(\mathbf{Q}) = d(\text{FULL})$ ,  $g(\mathbf{Q}) = g(\text{FULL})$ , and  $R(\mathbf{Q}) \leq 1 + \frac{f(P)}{P^t}$  for all sufficiently large  $P$ .*

For both beamforming and precoding strategies, we can achieve the full-CSIT diversity and array gains with a feedback rate of 1 bit per channel state asymptotically as  $P \rightarrow \infty$ . By precoding, we can approach this 1-bit rate asymptote with the

much faster  $\frac{\omega(1)}{P^t}$  decay compared to the  $\frac{\omega(1) \log P}{P}$  decay with beamforming. We now show that the  $1 + \frac{\omega(1)}{P^t}$  decay is the best possible up to  $o(1)$  multipliers in the term  $\frac{\omega(1)}{P^t}$ .

**Theorem 5.** *There is a constant  $C_3 > 0$  such that for any quantizer  $\mathbf{Q}$ , if  $d(\mathbf{Q}) = d(\text{FULL})$  and  $g(\mathbf{Q}) = g(\text{FULL})$ , then  $R(\mathbf{Q}) \geq 1 + \frac{C_3}{P^t}$  for all sufficiently large  $P$ .*

*Proof:* Let  $R = \frac{C_3}{P^t}$  with the constant  $C_3$  to be specified later on. Also, let  $\mathbf{Q} = \{\mathbf{X}_n, \mathcal{E}_n, \mathbf{b}_n\}_{\mathcal{I}}$  with  $|\mathcal{I}| \geq 3$ , and suppose that for every  $P_0 \in \mathbb{R}$ , there exists  $P \geq P_0$  such that  $R(\mathbf{Q}) < 1 + R$ . Using the same arguments as in the proof of Theorem 3, it is sufficient to show that  $\mathbf{Q}$  will not be able to achieve the full-CSIT diversity and array gains.

The properties of  $\mathbf{Q}$  imply the existence of some  $i \in \mathcal{I}$  with  $P(\mathbf{h} \in \mathcal{E}_i) \geq 1 - R$  (see the proof of Theorem 3). With  $f(\mathbf{h})$  denoting the probability density function of  $\mathbf{h}$ , we then have

$$\text{SER}_r(\mathbf{Q}) = \sum_{n \in \mathcal{I}} \int_{\mathcal{E}_n} Q(\sqrt{2\|\mathbf{X}_n \mathbf{h}\|^2 P/r}) f(\mathbf{h}) d\mathbf{h} \quad (38)$$

$$\geq \int_{\mathcal{E}_i} Q(\sqrt{2\|\mathbf{X}_i \mathbf{h}\|^2 P/r}) f(\mathbf{h}) d\mathbf{h} \quad (39)$$

$$= \int_{\mathbb{C}^T} Q(\sqrt{2\|\mathbf{X}_i \mathbf{h}\|^2 P/r}) f(\mathbf{h}) d\mathbf{h} - \quad (40)$$

$$\underbrace{\int_{\mathcal{E}_i^c} Q(\sqrt{2\|\mathbf{X}_i \mathbf{h}\|^2 P/r}) f(\mathbf{h}) d\mathbf{h}}_{\leq 1} \quad (41)$$

$$\geq \underbrace{\inf_{\mathbf{X}_i \in \mathcal{X}} \int_{\mathbb{C}^T} Q(\sqrt{2\|\mathbf{X}_i \mathbf{h}\|^2 P/r}) f(\mathbf{h}) d\mathbf{h}}_{\triangleq \text{SER}_r(\text{OPEN})} - \quad (42)$$

$$\int_{\mathcal{E}_i^c} f(\mathbf{h}) d\mathbf{h} \quad (43)$$

$$\geq \text{SER}_r(\text{OPEN}) - R \quad (44)$$

where  $\text{SER}_r(\text{OPEN})$  represents the open-loop SER with no feedback. Note that  $d(\text{OPEN}) = t$  but  $g(\text{OPEN}) < g(\text{FULL})$ . We thus choose the constant in the statement of the theorem as  $C_3 = \frac{1}{2}(1/g(\text{OPEN}) - 1/g(\text{FULL}))$ . Noting that  $\text{SER}_r(\text{OPEN}) \sim 1/g(\text{OPEN})P^{-t}$  and  $R(\mathbf{Q}) \leq C_3 P^{-t}$ , the lower bound in (44) then becomes  $\sim \frac{1}{2}(1/g(\text{OPEN}) + 1/g(\text{FULL}))P^{-t} > 1/g(\text{FULL})P^{-t}$ . Hence, either  $d(\mathbf{Q}) < d(\text{FULL}) = t$ , or if  $d(\mathbf{Q}) = t$ , we have  $g(\mathbf{Q}) < g(\text{FULL})$ . This concludes the proof. ■

We can therefore conclude that for precoding, the necessary and sufficient feedback rate that guarantees the full-CSIT gains is  $1 + \frac{\omega(1)}{P^t}$ , up to  $o(1)$  multipliers in the term  $\frac{\omega(1)}{P^t}$ .

#### ACKNOWLEDGEMENT

This work was supported in part by the NSF Award CCF-1218771.

#### REFERENCES

- [1] D. J. Love, R. W. Heath, Jr., V. K. N. Lau, D. Gesbert, B. D. Rao, and M. Andrews, "An overview of limited feedback in wireless communication systems," *IEEE J. Select. Areas Commun.*, vol. 26, no. 8, pp. 1341–1365, Oct. 2008.

- [2] D. J. Love, R. W. Heath, Jr., and T. Strohmer, "Grassmannian beamforming for multiple-input multiple-output wireless systems," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2735–2747, Oct. 2003.
- [3] S. Zhou, Z. Wang, and G. B. Giannakis, "Quantifying the power loss when transmit beamforming relies on finite-rate feedback," *IEEE Trans. Wireless Commun.*, vol. 4, no. 4, pp. 1948–1957, Jul. 2005.
- [4] V. Raghavan, R. W. Heath, Jr., and A. M. Sayeed, "Systematic codebook designs for quantized beamforming in correlated MIMO channels," *IEEE J. Select. Areas Commun.*, vol. 25, no. 7, pp. 1298–1310, Sept. 2006.
- [5] A. Narula, M. J. Lopez, M. D. Trott, and G. W. Wornell, "Efficient use of side information in multiple antenna data transmission over fading channels," *IEEE J. Select. Areas Commun.*, vol. 16, no. 8, pp. 1423–1436, Oct. 1998.
- [6] J. C. Roh and B. D. Rao, "Transmit beamforming in multiple antenna systems with finite rate feedback: A VQ-based approach," *IEEE Trans. Inf. Theory*, vol. 52, no. 3, pp. 1101–1112, Mar. 2006.
- [7] J. C. Roh and B. D. Rao, "Design and analysis of MIMO spatial multiplexing systems with quantized feedback", *IEEE Trans. Signal Process.*, vol. 54, no. 8, pp. 2874–2886, Aug. 2006.
- [8] V. K. N. Lau, Y. Liu, and T.-A. Chen, "On the design of MIMO block-fading channels with feedback-link capacity constraint," *IEEE Trans. Commun.*, vol. 52, no. 1, pp. 62–70, Jan. 2004.
- [9] J. Zheng, E. R. Duni, and B. D. Rao, "Analysis of multiple-antenna systems with finite-rate feedback using high-resolution quantization theory," *IEEE Trans. Signal Process.*, vol. 55, no. 4, pp. 1461–1476, Apr. 2007.
- [10] C. K. Au-Yeung and D. J. Love, "On the performance of random vector quantization limited feedback beamforming in a MISO system," *IEEE Trans. Wireless Commun.*, vol. 6, no. 2, pp. 458–462, Feb. 2007.
- [11] W. Santipach and K. Mamat, "Tree-Structured Random Vector Quantization for Limited-Feedback Wireless Channels," *IEEE Trans. Wireless Commun.*, vol. 10, no. 9, pp. 3012–3019, Sept. 2011.
- [12] K. K. Mukkavilli, A. Sabharwal, and E. Erkip, "On beamforming with finite-rate feedback for multiple antenna systems," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2562–2579, Oct. 2003.
- [13] P. Xia and G. B. Giannakis, "Design and analysis of transmit-beamforming based on limited-rate feedback," *IEEE Trans. Signal Process.*, vol. 54, no. 5, pp. 1853–1863, May 2006.
- [14] S. A. Jafar and S. Srinivasa, "On the optimality of beamforming with quantized feedback," *IEEE Trans. Commun.*, vol. 55, no. 12, pp. 2288–2302, Dec. 2007.
- [15] G. Jönsson, M. Skoglund, and B. Ottersten, "Combining beamforming and orthogonal space-time block coding," *IEEE Trans. Inf. Theory*, vol. 48, no. 3, pp. 611–627, Mar. 2002.
- [16] L. Liu and H. Jafarkhani, "Application of quasi-orthogonal space-time block codes in beamforming," *IEEE Trans. Signal Process.*, vol. 53, no. 1, pp. 54–63, Jan. 2005.
- [17] D. J. Love and R. W. Heath, Jr., "Limited feedback unitary precoding for orthogonal space-time block codes," *IEEE Trans. Signal Process.*, vol. 53, no. 1, pp. 64–73, Jan. 2005.
- [18] ——, "Limited feedback unitary precoding for spatial multiplexing," *IEEE Trans. Inf. Theory*, vol. 51, no. 8, pp. 2967–2976, Aug. 2005.
- [19] S. Zhou and B. Li, "BER criterion and codebook construction for finite rate precoded spatial multiplexing with linear receivers," *IEEE Trans. Signal Process.*, vol. 54, no. 5, pp. 1653–1665, May 2006.
- [20] S. Ekbati and H. Jafarkhani, "Combining beamforming and space-time coding using quantized feedback," *IEEE Trans. Wireless Commun.*, vol. 7, no. 3, pp. 898–908, Mar. 2008.
- [21] V. Tarokh, H. Jafarkhani, and A. R. Calderbank, "Space-time block codes from orthogonal designs," *IEEE Trans. Inf. Theory*, vol. 54, no. 5, pp. 1456–1467, July 1999.
- [22] E. Koyuncu and H. Jafarkhani, "Very low-rate variable-length channel quantization for minimum outage probability," *IEEE Data Compression Conf.*, Mar. 2013. [Online]. Available: <http://arxiv.org/abs/1210.8441>
- [23] ——, "Distributed beamforming in wireless multiuser relay-interference networks with quantized feedback," *IEEE Trans. Inf. Theory*, vol. 58, no. 7, pp. 4538–4576, July 2012.
- [24] ——, "A systematic distributed quantizer design method with an application to MIMO broadcast channels, *IEEE Data Compression Conf.*, Mar. 2010.
- [25] F. D. Côté, I. N. Psaromiligkos, and W. J. Gross, "A Chernoff-type lower bound for the Gaussian Q-function," Mar. 2012. [Online]. Available: <http://arxiv.org/pdf/1202.6483>
- [26] P. K. C. Wang, "On a class of optimization problems involving domain variations," *Intl. Symp. on New Trends in System Analysis*, Dec. 1976.